DOCUMENT RESUME

ED 152 813 95 TM 006 976

AUTHOR Smith, Jeffrey K.
TITLE Perspectives on Mastery Learning & Mastery Testing. ERIC/TM Report 63.
INSTITUTION ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princeton, N.J.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO ERIC-TM-63
PUB DATE 77
NOTE 29p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Academic Achievement; Cutting Scores; Early Childhood Education; Educational Strategies; Elementary Secondary Education; *Group Instruction; *Literature Reviews; *Mastery Learning; *Mastery Tests; Post Secondary Education; *Teaching Methods; Test Interpretation

ABSTRACT

Mastery learning is described as a system for providing a unit-by-unit sequence, and arranging that each unit be successfully mastered by each student before they begin the next unit. This process requires mastery tests to control the rate of student progress. The proper setting of score levels to represent mastery is a difficult task. The claim is made that, although group-based mastery learning may often delay the progress of the faster students at first, eventually the general result tends to be more efficient for learning for all students as compared to other group instructional processes. Bibliographical references are appended. (Author/CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED152813

ERIC/TM Report 63

PERSPECTIVES ON MASTERY LEARNING & MASTERY TESTING

by Jeffrey K. Smith

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM006 976

2

ERIC Clearinghouse on Tests, Measurement & Evaluation

The material in this publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Prior to publication, the manuscript was submitted to qualified professionals for critical review and determination of professional competence. This publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of either these reviewers or the National Institute of Education.

1977 ERIC/TM Report 63

PERSPECTIVES ON MASTERY LEARNING & MASTERY TESTING

by Jeffrey K. Smith

Cover Design: *Optical and Geometrical Patterns and Designs*
by Spyros Horemus, Dover Publications, Inc., 1970

Copyright © 1977 by ERIC Clearinghouse on Tests, Measurement, and Evaluation.
All rights reserved.

CONTENTS

Introduction 5

What Is Mastery Learning? 5

The Theory of Mastery Learning 6
- The Carroll Model 8
- The Concept of Aptitude 8

Procedures for Implementing Mastery Learning 11

Mastery Testing and Mastery Learning 13
- Distinguishing Mastery Tests from Other Criterion-Referenced Tests 14
- Research on Mastery Testing 15
- Construction of Mastery Tests 15
- Evaluation of Mastery Tests and Setting of Mastery Levels 15

Research on Mastery Learning 17
- Successful and Unsuccessful Mastery Learning Programs 17
- What Makes Mastery Learning Work? 18

Comments and Observations 20
- Criticisms of Mastery Learning Theory 20
- The Future of Mastery Learning 22
- A Final Word 23

References 23

INTRODUCTION

One of the basic problems of group instruction is that groups invariably consist of a collection of individuals. While groups are fine for such things as bridge games and debating societies, they present a problem for educators. The problem manifests itself in the following fashion. In a class of, say, 30 students, some will learn more quickly than the rest, some will learn at a moderate pace, and some will be slow. Even if all students enter the class with no prior knowledge of the subject matter, they will soon begin to differ with respect to their progress in the subject and become difficult to teach. Unfortunately, this phenomenon is not restricted to classes of 30; it works almost the same way with classes of 10, 20, or 50.

One method of addressing the problems of group instruction is the mastery learning theory developed by Benjamin Bloom and his students (11, 17, 19). In his 1968 work, "Learning for Mastery," Bloom outlined the basic tenets of the theory. Since the appearance of that paper, over 100 others on mastery learning have been published in journals or presented at conferences (72), and over 25 dissertations have investigated mastery learning strategies. Meanwhile, mastery learning strategies have been developed in disciplines ranging from physics (48) to physical education (5) throughout the United States and in many foreign countries (72).

In addition to gaining popularity, mastery learning has gained a variety of definitions. In its most general usage, mastery learning can refer to any instructional strategy that requires a student to display expertise at a predetermined level; in its most specific usage, it refers to the set of instructional strategies developed by Bloom. While I will consider variations of the former usage in this paper, my primary focus will be on Bloom's strategies.

The purpose here is to present mastery learning theory for those who have not looked at it closely before. Following a definition and a brief look at the various usages of the term, the theoretical framework for mastery learning will be presented along with operating procedures for implementing mastery learning in a classroom. The emerging field of mastery testing will be examined, and the research on mastery learning will be summarized. Finally, some of the implications of mastery learning for educational philosophy and practice will be discussed.

WHAT IS MASTERY LEARNING?

The question posed in the heading grows more difficult to answer almost daily. In "Learning for Mastery," Bloom outlined a set of instructional strategies for teaching students in classroom settings. Very briefly, Bloom included the following strategies:

1. The units or chapters in the course being taught are analyzed in order to develop tests for assessing student progress on a unit-by-unit basis (called *formative* tests).
2. All students receive instruction on the initial unit.
3. All students are administered the formative test for that unit.
4. All students are informed of their specific weaknesses with respect to the unit and are provided with alternative learning materials and assistance in order to reach a predetermined level of mastery.*
5. The process begins again with the next unit.
6. At the end of the course, all students are graded by means of a test that covers the

*A discussion of what "predetermined mastery levels" are and how to get them will be presented later.

entire course and has fixed cutoffs for all grades. This kind of test is called a criterion-referenced summative test.**

Details of these strategies will be presented in a later section.

It is important to note that in Bloom's development of mastery learning, the students remain together in their instruction; in fact, instruction is initially presented in a traditional group approach. It is the remediation of weaknesses that is individualized.

There have been a number of variations of this type of instruction which have been called mastery learning approaches. Variations usually involve elimination of whole-group instruction in favor of individualized instruction. Perhaps the best known of these systems is the Keller Personalized System of Instruction (PSI) (46). Under Keller's system, students work individually on units, taking unit tests when they believe they are ready. They must reach a certain level of mastery on the test in order to move to the next unit. Final grades are based upon the number of units that are successfully completed. Under the Keller system, there actually need not be any instruction by the teacher in the traditional sense, beyond helping students individually.*

Other modifications include testing on skills prior to instruction so that students who already possess skills need not spend time on them. Various computer-assisted instructional programs and other programmed materials have also fallen under the rubric of mastery learning from time to time.

What is common to all strategies called mastery learning is the requirement that students display a particular level of expertise on a test of a skill or a unit before proceeding to the next skill or unit. Beyond that commonality, the systems vary greatly.

This variety leaves those who wish to use the term in something of an orthodoxy conflict. Using the term to refer just to the strategies developed by Bloom is somewhat limiting and it is not even used by Bloom in that fashion (16). On the other hand, if the term is used for any strategy that fits the description of the preceding paragraph, then mastery learning can simply mean highly individualized instruction or instruction that is group based.

Since the term is used to refer to a variety of instructional methods, one should be careful to find out what an author who is using the term is writing about. In this paper we will be primarily concerned with Bloom's work and will use *mastery learning* to refer to Bloom's strategies and *individualized mastery learning* to refer to other strategies.

THE THEORY OF MASTERY LEARNING

As children progress through school, the discrepancy between the students who perform at the higher levels and those who perform at the lower levels grows wider and wider. Figure 1 displays this discrepancy in reading comprehension for students in a large metropolitan area (66).

It can be seen that the differences even at age seven are considerable and that by age twelve the discrepancy between the tenth and ninetieth percentiles is almost five years! While there are many reasons offered for the increasing discrepancy between highest and lowest, most reasons have something to do with the cumulative effect of

---

**Summative tests "sum up" a person's progress. They are used for grading or certification. Formative tests are used to make educational decisions about a person in a more immediate sense. These results are used to guide instruction for the individual.

*For a comparison of the Keller system with the Bloom system, see Block, J.H., *Schools, Society, and Mastery Learning*, pp. 20-26.

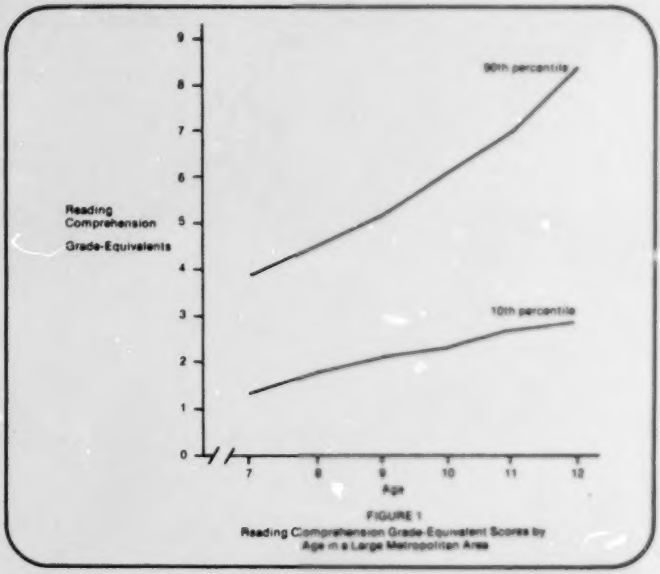

FIGURE 1
Reading Comprehension Grade-Equivalent Scores by Age in a Large Metropolitan Area

differences in aptitude over the course of the school years. Bloom (16) contends that this need not be the case, and in fact, with appropriate learning conditions, the situation in Figure 1 could be made to resemble the hypothetical situation in Figure 2.

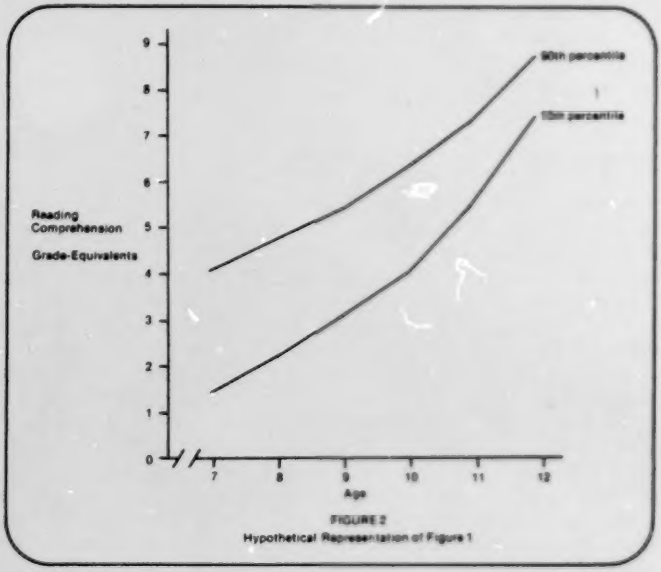

FIGURE 2
Hypothetical Representation of Figure 1

Bloom believes that the situation in Figure 1 is not the inevitable result of individual differences, but rather is *caused* by the educational system.

In *Human Characteristics and School Learning* (16), Bloom summarizes much of the work that has been done on mastery learning since his "Learning for Mastery" appeared in 1968 and incorporates the findings in a more general theory of school learning than was presented in that paper. He states that: "Most students become very similar with regard to learning ability, rate of learning, and motivation for further learning—when provided with favorable learning conditions."

Since it was the mastery learning research that Bloom in large part was using as evidence for this statement, an investigation into the theory seems worthwhile.

The Carroll Model

Perhaps the single most influential work on the development of mastery learning theory is John Carroll's 1963 article, "A Model of School Learning" (22). In the article Carroll examines the influences on the learning of a single instructional task by an individual. Carroll postulates that the degree of learning on the task is a function of two time variables—time spent in learning and time needed for learning (see Figure 3 on page 9).

$$\text{degree of learning} = f\left(\frac{\text{time spent}}{\text{time needed}}\right)$$

The equation simply states that if the time spent in learning equals the time needed, then learning will be complete. If time spent is less than time needed, then learning will be incomplete. Carroll says that time needed in learning is influenced by the aptitude of the learner, his ability to understand the instruction being presented, and the quality of instruction being presented. Further, Carroll says that the time spent in learning is determined by the time allowed for learning (opportunity to learn) and by the perseverance of the learner. These variables might be displayed in an "influence diagram" such as Figure 3.

The Concept of Aptitude

The most interesting aspect of this model, and certainly the one with the most relevance for mastery learning theory, is Carroll's view of aptitude. He suggests that aptitude be thought of as the amount of time needed in learning for an individual when the quality of instruction is optimal for that individual and when he is persevering on the task. The author notes that aptitude is specific to the task at hand and is subject to two influences: traits or characteristics of learners that may be genetically determined or may be based on "generalized prior learnings" and more specific learnings that are relevant to the present task. Carroll states that even when the quality of instruction is optimal for all students and when all students perfectly understand the presentation of instruction, there is considerable variance among students with respect to aptitude. Using studies by Glaser (34) and Atkinson (7), Bloom (16) estimates that a student at the tenth percentile in aptitude for a task may take five or six times longer than a student at the ninetieth percentile.

If, however, students could be made to be very similar with respect to the specific prior learnings relevant to a task, then the differences in time needed on the task could be diminished. The degree of diminution would depend on the importance of the spe-

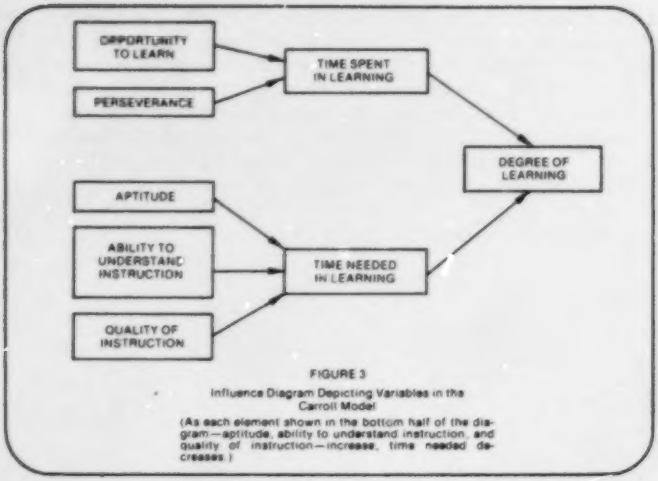

FIGURE 3

Influence Diagram Depicting Variables in the Carroll Model

(As each element shown in the bottom half of the diagram—aptitude, ability to understand instruction, and quality of instruction—increase, time needed decreases.)

cific learnings in comparison to the generalized prior learnings or genetic influences. This possibility of making students similar is crucial to mastery learning theory. Two basic tenets of mastery learning theory are:

1. Students *can* be made very similar with respect to specific prior learnings relevant to a task.
2. As one progresses through a course, the strongest influence on time needed for a new unit is achievement on prior units.

Therefore, if all students master Unit 1 before proceeding to Unit 2, then the variance in time needed among students for Unit 2 will be diminished. If this procedure is followed through the course, then by, say, Unit 10 students would be much more similar with respect to time needed to master that unit than they were on the earlier units. At this point an illustration may be useful.

Consider two high school introductory algebra courses. Mr. Daly teaches algebra in a traditional fashion. Mr. West uses a mastery learning approach. At the end of the first two weeks of school, both classes have finished instruction on Unit 1 and a unit test is given covering the essential features of that unit. Both classes do fairly well on the test with about 70 percent of the students in each class scoring above 85 percent correct. Mr. Daly proceeds to Unit 2. Mr. West devotes three extra days of instruction to Unit 1 with specific homework assignments and peer tutoring designed to bring the remaining 30 percent of the class up to an 85 percent correct minimum.

Instruction on Unit 2 also takes two weeks for each class (with Mr. West's class being three days behind). During instruction on Unit 2, however, Mr. Daly needs to spend some time clearing up problems for students who didn't master Unit 1. These students are at a disadvantage on Unit 2 because it is related to Unit 1. Also, one or two other students may fail to master Unit 2 completely because of instruction moving too quickly, inattentiveness, absence, and so forth. Mr. Daly's class is beginning to spread out in achievement, but they are covering one unit every two weeks.

Mr. West on the other hand, has a little less trouble with Unit 2 since everyone learned Unit 1. Still, after the unit test is given, a few extra days may be needed for review, and Mr. West takes them.

At the end of ten weeks, Mr. Daly has covered five units. The distribution of achievement in his class on these units is depicted in the upper graph of Figure 4. Mr. West has covered only four units; his class has fallen a full unit behind. The distribution of achievement in his class on those units, however, is depicted in the lower graph of Figure 4. Additionally, Mr. West has found that he can present instruction initially in nine days instead of ten, and that remediation only takes a day or two. Mr. West's students, almost all of whom have experienced continued success with algebra, show a general positive attitude toward the subject.

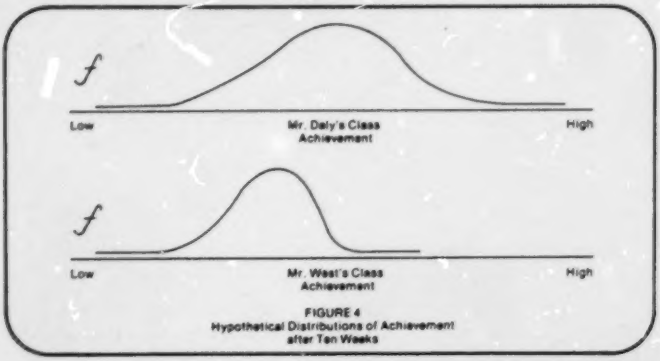

FIGURE 4
Hypothetical Distributions of Achievement after Ten Weeks

Mr. Daly's slower students are getting increasingly frustrated with algebra instruction and increasingly adept at providing diversions for their classmates during class. The faster students no longer need two weeks of instruction in order to master a unit and sometimes find algebra boring.

Near the end of the semester, Mr. Daly devotes two weeks to review of the units covered up to that point. Since students need help on a variety of concepts, the review is necessarily spotty. Mr. West needs most of the last two weeks to arrive at the same unit as Mr. Daly, but he finds very little review necessary. In essence, he spent his allocated two weeks of review time throughout the semester, helping students whenever necessary. Figure 5 depicts the achievement of the two classes over the course of the semester.

This example clearly portrays the mastery learning approach in a most favorable light. The extent to which results such as depicted for Mr. West are actually obtained will be examined later. Right now, let's look at Mr. West and Mr. Daly with respect to the variables in the Carroll model.

First, there is the opportunity to learn. If a student needs twelve days of instruction in order to learn a task and is only allowed ten, then learning cannot be complete for that student. In mastery learning, instead of fixing the amount of time for a task and letting the degree of learning vary, the degree of learning is fixed and the time is allowed to vary. In the beginning of a course, this means that some students will not progress as fast as they might otherwise since extra time is allocated for slower students.

Second, there is the ability to understand instruction and the quality of instruction. These variables interact and determine the *efficiency* of instruction (22). In the traditional approach, after a few units, some students can no longer understand instruction since they have not mastered prior concepts being used as building blocks of new concepts. One of the goals of mastery learning is to insure that students have mastered prior concepts and are ready for new concepts.

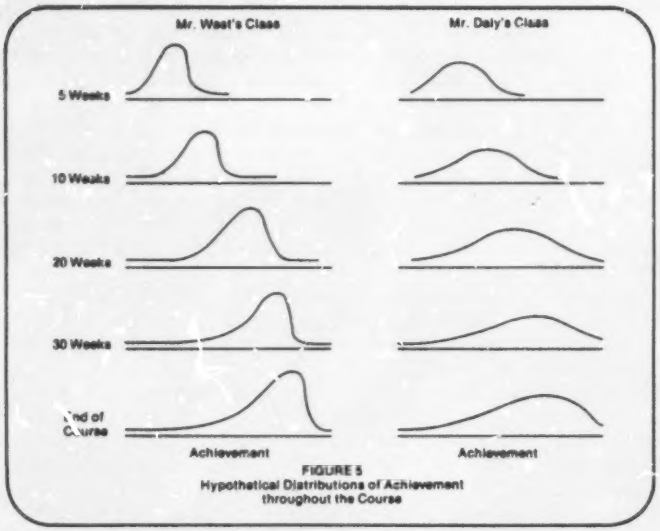

FIGURE 5
Hypothetical Distributions of Achievement throughout the Course

Third, there is perseverance. Few human beings will persist at anything if they repeatedly fail at it. In mastery learning, students learn that the teacher will stay with them and provide help until they *do* learn. Students are provided with frequent and regular successes in mastery learning, and they develop positive attitudes toward the subject, the school, and themselves as learners.

Finally, there is aptitude. Aptitude in this model is composed of specific prior learnings, generalized prior learnings, and genetic factors. In a traditional approach, the time needed to learn grows increasingly variant among students since specific prior learnings and generalized prior learnings grow more divergent. In mastery learning, since students reach a predetermined level of mastery before proceeding, they are kept similar with respect to specific learnings and the time-needed variation is cut back. Further, the generalized prior learnings gradually become more similar. Thus, the time-needed variation is diminished further. How similar students become with respect to time needed in learning is a question that has not been resolved.

The foregoing discussion presents the essence of the theory behind mastery learning. At this point it should be apparent how different Bloom's approach is from individualized approaches. In a later section, we will see how much of the theory has been verified by research findings and in what fields mastery learning works best.

PROCEDURES FOR IMPLEMENTING MASTERY LEARNING

It should be noted at the outset of this section that mastery learning is not a curriculum theory or a theory that specifies what *should* be taught. It is instead a theory on how to *manage* instruction in a given curriculum.

Implementation of a mastery-learning strategy begins with a specification of the behavioral objectives of a course. Bloom, Hastings, and Madaus (19) provide a detailed discussion of how this might be accomplished. This list of objectives provides a statement of what the course is about. This statement should be specific enough to be the basis for writing a summative (final) exam.

Once the objectives are written, we turn to the organization of curricular materials. Most published textbook series or sets of instructional materials are organized into chapters or units. (Even when materials are prepared by the teacher, they frequently have a kind of chapter or unit organization.) These chapters or units are typically arranged sequentially, to the extent that they can be, so that learning in the course is a cumulative process. Each unit must be thoroughly analyzed in such a fashion that the important terms, facts, concepts, principles, and so forth, are extracted. After this has been done, a formative test based on this analysis can be constructed for the unit. This procedure alone may be quite beneficial. Teachers often discover that much of what is covered in class is not really essential to mastery of the course objectives; while some of this content may be related to affective outcomes of instruction, much of it is discarded after such an analysis.

Once the formative tests have been constructed, instruction can begin. Instruction is initially presented to students in the same fashion that it would be under a nonmastery approach. At the completion of instruction on the first unit, the first formative test is given. This test serves two purposes. First, it lets the *teacher* know how well the class did in general, what aspects of instruction need to be reviewed with the entire class, and what aspects individual students need to review. Second, it provides the *learner* with feedback on his individual progress and needs.

After the administration of the formative test, a review period begins. This is somewhat different than under traditional instruction where review typically occurs *prior* to a unit test. In the review period, students who need help are provided with alternative learning materials and assistance in mastering aspects of the unit that were not mastered originally. In high school or college courses, this review can occur outside of scheduled class time. It might also include assignment of peer tutors. Under mastery learning procedures, students are encouraged to work on weaknesses in groups. These groups can be formed on the basis of similar weaknesses (so that all can work on the same aspect of the unit), different weaknesses (so that students can help each other), or affective characteristics of learners (so that students are grouped with peers with whom they will work productively). These groups will naturally change from unit to unit as students have different combinations of needs. The concept of a slow group and a fast group doesn't exist in mastery learning.

The number and nature of the groups will vary according to grade and subject matter. Elementary teachers will probably find it taxing to manage more than two groups. It is helpful if teachers can individualize the review (68). Also, at the high school and college levels, groups may meet outside of class.

When the student believes he has mastered the material that was not learned initially, he takes an alternate form of the formative test to check for mastery. In some situations, teachers may want to retest only the material with which the student originally had difficulty. Later, we will discuss what percent correct on the formative test is the most desirable level of mastery. For now, it would be useful to think of about 80 to 90 percent as a desirable level of mastery.

The review period continues until all or almost all students have reached a desirable level of mastery (perhaps 95 percent). Making a decision as to when to stop the review and begin the next unit is a difficult process. The teacher should consider the following factors in deciding when to move on:

1. When the class moves to the next unit, students who have not mastered the unit are going to have difficulty with succeeding units.* Therefore, especially early in a course, additional time and help for a few will benefit *all* later (as in the example

*If students seem to be *totally unable* to learn an early unit, perhaps they should not be in the course, particularly at the high school and college level.

with Mr. West and Mr. Daly).

2. If the test shows that mastery has not been achieved, but informal evaluation (such as individualized questioning) indicates mastery of the required skills, the teacher should check to make sure that the test itself is a reliable measure.
3. The teacher may also at this point question the need for that particular unit in the overall scheme of the course. If the unit seems less than critical, the teacher may want to proceed with fewer students at mastery level than usual.

Of fundamental importance, however, is the teacher's sincere and patient attempt to bring all students to mastery. It is inherent in the theory that if such an attempt is made, it will be successful.

Once the decision to move to the next unit has been made, the process simply begins again with the next unit. Grading is based solely on a summative examination given at the end of the course. The formative tests do not contribute to the final grade. There are two reasons for this. First, this practice eliminates anxiety on the part of the students as they approach the formative tests, which helps make the tests a true learning tool. Second, since all or most students will reach mastery sooner or later, counting the formative test would simply be rewarding the faster students. Grades are assigned in a criterion-referenced fashion. Anyone scoring over a certain percentage correct on the final examination gets an "A." The criterion for setting the "A" level should be whatever it would have been under traditional instruction.

The foregoing discussion covers the basic features of implementing mastery learning procedures in the classroom. But before attempting such an implementation, the author recommends that the implementor read *Mastery Learning in Classroom Instruction* (13).

MASTERY TESTING AND MASTERY LEARNING

In the individualized systems for mastery learning as well as in Bloom's strategies, one area of interest for research and practice is testing for mastery. Mastery testing has become a field unto itself. Unfortunately, like mastery learning, the term *mastery testing* has acquired a variety of meanings. We will attempt to provide a general delineation of the term here as well as an overview of the area.

Basically, mastery testing is a subcategory under the more general heading of criterion-referenced testing. Criterion-referenced testing is a fairly recent concept in test theory, first appearing under that nomenclature in 1963 (35). The underlying concept of criterion-referenced testing is that a person's performance on a test ought to be related to (referenced to) a criterion of performance (for example, "drives well enough to have a license," "can throw a ball over 100 feet," "reads well enough to comprehend a newspaper article") rather than to the achievement of a set of arbitrarily chosen comparable people taking the same test (for example, "eighty-sixth percentile on the college boards," an "I.Q. of 116," "below grade level in reading").

The difference between criterion-referenced testing and norm-referenced testing is not necessarily in the test but in the *interpretation* of the test results. The emphasis is on the word *referenced.* Consider, for example, a driver's license examination. If one norm referenced the results of a driving test, one could make a statement such as the following: "John's driving ability is at the thirty-sixth percentile of all men, age 16 to 21." This is useful information for a variety of different purposes, but it does not tell one whether John should have a license or not. If one criterion referenced the results, one could say: "John's driving ability is such that in all likelihood he would not be a hazard when driving on public thoroughfares." For licensing purposes, that is the information that is needed. The point is that the test could be the same test in both in-

stances. There are, of course, implications for test construction if one approaches construction from a criterion-referenced as opposed to a norm-referenced perspective, but the essence of the difference is in the interpretation of the results.

Distinguishing Mastery Tests from Other Criterion-Referenced Tests

The notion of criterion referencing is quite compatible with mastery learning. In mastery learning one is not interested in comparing individuals' test scores. One is concerned with how well students are performing with respect to the content of the course (in both group-based and individualized mastery learning). However, the evaluation need in mastery learning is a little more specific than simply referring a person's score to various levels of a criterion. The need is to make instructional decisions about students based upon their scores.

Essentially, the decision revolves around the question, "Has this student mastered the content of the unit?" Refined to its implications for testing, the question becomes "What level of performance on this test will we accept as sufficient evidence of mastery of the unit?" The evaluation here is an integral part of the instructional process. The process could not continue as designed without it. Using the criterion-referenced test in a formative manner as part of the instructional process is one of the distinguishing characteristics of a mastery test.

Whether one is testing a single skill or a chapter of a textbook, the possible outcomes of the testing situation may be viewed from two perspectives. First, ability and hence performance may be thought of as distributed in some fashion on a continuum. This might be more easily envisioned under the Bloom-type construction where students have mastered some aspects of a unit and not others. It is also possible, however, to imagine individuals with varying degrees of mastery of a single skill.

The second perspective on this issue is that ability is dichotomous with respect to the test. A person either has a skill or not—"knows" the unit or doesn't know it. There are consequences for test evaluation and for the setting of mastery levels for both perspectives.

The term *mastery test* connotes a high ability in performance on a test, a "mastery" of the content. This feature of mastery tests also distinguishes it from the more general criterion-referenced tests. The rationale behind requiring a high level of mastery is that with the measurement of single skills or concepts, students either have mastered the skill (and will get all or nearly all items right) or they have not (and will get few items right). This high level of performance need not be necessary in mastery tests, but it is a characteristic frequently associated with them.

A final consideration in identifying a criterion-referenced test as a mastery test is the scope and length of the test. Mastery tests tend to measure a limited number of concepts, skills, or objectives. Criterion-referenced tests, on the other hand, are as unlimited in scope and length as norm-referenced tests.

Thus, there are three possible distinguishing characteristics of the subcategory of mastery tests:

1. Educational use of results usually formative
2. Fixed, typically (though not necessarily) high level of acceptable performance, indicating "mastery" of the content being measured
3. Limited scope and length, often measuring a single skill, concept, or unit in a course

If one combines these three characteristics, one obtains a test that would provide an instructor with the basis for deciding whether a student should work more on the present skill (unit, concept) or proceed to the next. That is what mastery tests are.

Research on Mastery Testing

Mastery tests are by nature criterion-referenced tests, and much of the research on criterion-referenced testing is also applicable to mastery testing. The research in this area focuses on three particular aspects:

1. Construction of mastery tests
2. Evaluation of the tests (in terms of validity and reliability)
3. Setting of performance levels acceptable for establishing "mastery"

Construction of Mastery Tests

Of the three aspects of mastery testing mentioned above, perhaps the one that is least well-defined by researchers is the construction of mastery tests. The difficulty arises in trying to establish procedures that are generalizable. The most thorough discussion of procedures for construction is the *Handbook on Formative and Summative Evaluation of Student Learning* (19). Bloom approaches the construction of mastery tests ("formative tests" in his terminology) from an analysis of the unit of instruction. Bloom recommends the construction of a "table of specifications" of the unit from which the mastery test can be developed. All topics in this table should be represented by items in the test. The table should contain only those aspects of the unit which are important for later learning in the course. Airasian (3) and Bosshart (20) have investigated the use of this procedure and have found that there are varying degrees of success with it.

Other work in this area by Reichman and Oosterhof (61), Emrick (29), and Hambleton (37) has typically involved test construction related to a single skill or objective. In the construction of tests of this type, precision in the specification of the objective being tested is directly related to the ease and clarity with which the items can be written. For a general discussion of this type of procedure, see Mayo (52).

Evaluation of Mastery Tests and Setting of Mastery Levels

The evaluation of mastery tests (and, more generally, criterion-referenced tests) has spawned an entirely new area of psychometrics. There are several problems encountered in evaluation of criterion-referenced testing that are not encountered in norm-referenced testing. These problems stem from differences in the theories underlying the two approaches. In norm-referenced testing, of particular interest are the differences among individuals with respect to the content being measured. Differences are assumed to exist, and to exist in a particular fashion (usually normally distributed). Thus, the test can be evaluated in terms of how well these differences are captured by the test. Items can be assessed by the degree to which they provide a basis for discriminating among individuals.

In criterion-referenced testing, the interest is no longer in making comparisons among individuals. Hence, much of what is important in norm-referenced testing is irrelevant to criterion-referenced testing (and thus to mastery testing).

Assumptions about the distribution of ability with respect to the content being measured are more difficult to make in criterion-referenced testing. What is of importance in criterion-referenced testing is the proper classification of individuals into mastery or nonmastery categories. This brings us to the third area of research on mastery testing: the determination of cut-off or mastery scores.

Research in these two areas (evaluation and determination of cut-off scores) has grown considerably over the last five to ten years. The suggestions made for this evalu-

ation and determination generally fall under one of three categories:

1. Approximations to existing test models: Research in this area involves using either classical test theory approaches (8, 25, 31) or item characteristic curve models (36).*
2. Development of new approaches and indices of reliability and validity: Most of the research in criterion-referenced testing is concerned with the development of new approaches to estimating reliability and validity. These usually involve the development of a new index of reliability (30, 41, 58, 63, 70) or a procedure for estimating the proportion of correct classifications into mastery and nonmastery categories (1, 29, 39, 49, 55, 69).
3. Nontechnical approaches: Some researchers have approached the problem from a perspective other than technical. Ebel (28) suggested that cut-off scores be determined from judges' ratings of item difficulty and relevance. Crambert (24) suggested that careful attention to test development standards is essential to assurance of validity in criterion-referenced testing. Block (12) recommends that the establishment of a cut-off score be related to the use of that score in the educational process. He suggests that the affective consequences of testing should be considered in setting cut-off scores. His research showed that varying the cut-off level affected both the cognitive and affective growth of students with respect to the subject. For a thorough review of many of the approaches presented here and others, see Meskauskus (54).

Cut-off scores for mastery tests are often in the range of 80 percent to 100 percent correct.

The position taken here is that it is necessary to evaluate the test in terms of its relationship to the instructional process in order to determine a cut-off score. The goal is to find a minimal score necessary to insure success on later units. This score will depend upon how important the content of the unit is and how well that content is represented by the test. It may be that the content of the unit is difficult to measure with the available measurement techniques. It may also be that the unit is not really critical to later learning but desirable if attainable. These possibilities should be considered in determining a cut-off score. Thus, if one combined Ebel's notion of assessing the difficulty of the items and their relevance to the instructional process with Block's concern for the affective consequences of the testing procedure, one should be able to determine a reasonable level of mastery for the mastery test. It is clear that with these concerns, one will establish different levels for different tests.

It should be noted that setting mastery cut-off levels is not a simple procedure and that the methods mentioned under categories 1 and 2 above are not realistic for the classroom teacher.

Before leaving the topic of mastery testing, it should be pointed out that test items almost always measure skills in addition to the skill under consideration. This idea is similar to Ebel's concept of difficulty and relevance, but the focus is on *what else* the student has to do in order to respond correctly even if he has the desired skill. For example, reading is involved in most items, even though the skill in question may be unrelated to reading. The additional skills necessary to respond successfully to an item above and beyond the skill in question might be thought of as "item baggage." When the test constructor is determining a cut-off score for a mastery test, he should always consider the ease with which a student can arrive at the wrong answer while possessing the skill of interest. It may be that the amount of item baggage is greater than

*Classical models and item characteristic curve models were developed for norm-referenced testing. It is not necessary here to explain them; it is only worth noting that both models operate under assumptions that in some respects are not applicable to criterion-referenced testing.

appears to the test constructor. A clinical evaluation of the test with students whose ability on the skill has been unambiguously determined a priori is recommended as a method of assessing item baggage. The amount of this undesirable entity on a test should be a consideration in determining a cut-off score. Stating a single percent correct for tests measuring different skills is very likely an unproductive endeavor. Since the difficulties of items do not always solely reflect difficulty inherent in the skill being tested, universal cut-off scores are rarely appropriate.

Setting cut-off scores requires teachers to make subjective decisions. It is a difficult process and an important one to classroom learning under mastery learning conditions (group or individualized).

RESEARCH ON MASTERY LEARNING

The research on mastery learning is usually based on one of two questions: Does it work? or What makes it work? We will look at both approaches here. The purpose of this section is not to provide an exhaustive review of the literature in mastery learning theory. For such a review, the reader is referred to Block and Burns (14). The purpose here is to look at selected literature that details some of the areas in which mastery learning has been tried and examine research that asks why it works and how it might be improved.

Successful and Unsuccessful Mastery Learning Programs

The claims made for mastery learning are far from modest. Keller's (46) seminal work is rather brazenly entitled "Goodbye Teacher. . ." and Bloom suggests that 90 percent of all students can learn at levels presently reserved for the top 10 percent. But claims are not results. In this section we will look at the results that have accumulated thus far. Once again, the focus is on the group-based mastery learning strategies.

Essentially, the research shows that successes have been found in a variety of subjects at a variety of age levels. Bloom (16) cites several studies in which mastery learning groups were compared to control groups:

1. Lee, *et al.*, (50) found that in Korea fifth and sixth grade students performed at a much higher level in arithmetic and science under mastery learning conditions than under control (more traditional) conditions. This study was carried out with a very large sample across a number of schools.
2. Pillet (59) also reports positive results in a mastery as compared to a control experiment in French as a foreign language with high school freshmen.
3. Kersh (47) in a study on fifth grade math and Jones, *et al.* (44), in a study on junior college business, economics, and biology courses also found mastery conditions to be superior to control conditions.

Block and Burns (14) report studies using a variety of experimental designs:

1. Anderson (4) and Block (11, 12), in courses in matrix algebra at the eighth grade and college levels, found mastery strategies superior to controls.
2. Wentling (71) found mastery groups superior to control groups in a high school auto mechanics course.
3. Fiel and Okey (32), in a study of a one-week course in graphs at the eighth grade level, report mastery groups performed better than groups under control conditions.

Studies showing differences in favor of mastery learning have been done on college-level courses in educational psychology (33, 38), statistics (53), test theory (2), and

algebra and English (64). At the elementary level, in fourth through seventh grade math, Collins (23) found mastery learning superior, and in an elementary level geometry study, Burrows and Okey (21) found low-aptitude fourth grade pupils in a mastery learning classroom performing as well as high-aptitude fifth graders under more traditional instruction.

While all of these studies report findings in favor of mastery learning (at varying levels of significance), only the Burrows and Okey study really show the magnitude of gains suggested by the theory. On the negative side, Davis (27) reports no significant differences between a mastery and control group in a college freshman English course, and Myers (56) reports similar results in a course in college geography. Furthermore, it is very likely that other studies of unsuccessful programs have not been published or presented. One may feel that the final tally is that many studies show success, but rarely to the levels predicted, and some show no success at all. However, reinspection of the results with an eye toward mastery learning theory leads to a somewhat more clear-cut picture.

Most of the subject areas in which mastery learning has been successful have some common characteristics. Typically, they are subjects in which there is a natural sequential ordering of the units (such subjects as math, science, and foreign language). One does not see a number of successes in subjects such as English, composition, social studies, literature, reading, and so forth. These are areas in which the content does not lend itself as well to a sequential ordering. Furthermore, courses for which mastery strategies work best typically involve students who have the basic cognitive prerequisites for the course, but little or no prior knowledge of the subject matter. This brings us to the second area of research—investigation into what makes mastery learning work.

What Makes Mastery Learning Work?

There is a fairly substantial body of literature on mastery learning that does not ask Does it work?, but rather addresses somewhat more refined questions such as:

1. Under what conditions does mastery learning work best?
2. What variables might affect outcomes under mastery strategies?
3. What changes occur in students and teachers when they work with mastery learning?

As mentioned above, mastery strategies have proven to work best with courses that are sequential in nature, with students who have the cognitive history to do well, and with all students starting on fairly equal footing. This naturally suggests courses in mathematics, science, foreign language, and training courses. According to Bloom (17), these courses might be thought of as emphasizing "convergent" thinking. Block and Burns (14), however, report positive findings in areas where more "divergent" thinking is required (historiography, geography, and so forth). Additionally, research in elementary reading instruction has been most encouraging (45, 67, 68).

So, while it is clear that sequentially ordered courses involving convergent thinking have enjoyed the greatest success under mastery learning conditions, it appears that courses that do not meet those specifications might also benefit from mastery learning strategies. In order to get a feeling for why mastery learning strategies might work for courses other than the "sequential/convergent" ones, we need to look at the variables that affect the success of mastery learning strategies.

Probably the single most important factor contributing to mastery learning successes is the requirement that all students master prior units before moving on to more diffi-

cult ones. This is the hallmark of both the Keller individualized system and the group-based strategies. This procedure ensures that students are not working on materials too difficult to be mastered by them at that time. Block and Burns (14) emphasize the importance of requiring unit mastery and cite several studies that support this position—for example, Anderson (4), Block (12), and Johnston and O'Neill (42).

Requiring unit mastery probably affects outcomes in more ways than simply ensuring that students are always working on materials at an appropriate difficulty level. Under this requirement, students are continually experiencing success in the course. This has an affective impact on students in addition to a cognitive one. In an attempt to find an aptitude treatment interaction between locus of control* and mastery learning, Reynolds and Gentile (62) found that "subjects overwhelmingly preferred the mastery learning procedures to traditional procedures." This is not surprising; people like to succeed. Studies by Levin (51), Ozcelik (57), Arlin (6), and Lee *et al.* (50) all show more positive attitudes under mastery conditions. This change toward more positive attitudes is not limited to the sequential/convergent subjects; it generalizes to any subject in which continued success can be experienced.

Katims (45) speculates that the primary reason that success has been found with mastery learning strategies in reading is that the strategies provide a framework for high quality instruction. He states:

> The stability provided by the regularity of the format (teach-practice-test-remediate) allows students to encounter written language in different forms (sentences, paragraphs, stories, poems, etc.) and successfully master a wide variety of language skills (e.g. making syllables, using a dictionary, identifying the main idea). Furthermore, with the extensive practice and testing provided by the model, both the teacher and the student are able to see the student acquire new skills. This gives confidence to both. Thus, the program is based on the premise that the mastery learning model *can* work in reading, by 1) providing verbal experience to students in a variety of ways, and in proportion to their need, and 2) by allowing both teachers and students to reap the affective benefits of regular success experiences.

Thus, there are three variables that might contribute to the success of a mastery learning strategy:

1. Requirement of unit mastery: This is probably the most important variable—certainly for the sequential/convergent courses. It is also important for other courses, for certainly in almost any course, there is an accumulation of knowledge and skills that aid in learning new skills, and *some* convergent thinking is required in almost all courses.
2. Affective benefits of instruction: As students succeed, their self-confidence increases, their attitude toward the subject improves, their interest in it increases, and their willingness to persevere in the next units increases.
3. Quality of instruction: The management of instruction (teach-practice-evaluate-remediate)—the notion of students working together in groups on problems, and the idea of students being informed of what and how they are to learn and how well they are progressing—combine to form a high level of instructional quality that would probably benefit students even without the aid of 1 and 2 above.

We have looked at the situations under which mastery learning works best and variables that might affect outcomes under mastery learning strategies. Now let us look at changes in students and teachers that occur under mastery strategies. Here we

*Locus of control is a psychological concept that concerns whether people feel that they are in control of their lives or are subject to external influences.

will consider findings that are critical of mastery learning as well as ones that are supportive.

Bloom (16) contends that mastery learning strategies improve the student's attitude toward the subject and the school, and improve his or her self-concept as a learner. Block and Burns (14) report a number of studies showing general positive changes in students after mastery instruction. Katims *et al.* (45) report enthusiasm among elementary school teachers using mastery learning in reading instruction.

There are, however, some negative findings and some general negative contentions in the literature. Block and Tierney (15) found mastery learning groups less interested in the subject matter after instruction, and Poggio *et al.* (60) found mastery learning students showing significantly greater anxiety about the final exam than students under control conditions. Both sets of authors are mastery learning proponents.

The strongest and most consistent attack upon group-based mastery learning is centered on what happens with the faster students. Critics of mastery learning such as Skaalvik (65) and Jones (43) contend that mastery learning retards the progress of the high aptitude student by continually requiring that he wait for his slower classmates. (In the PSI method, since students work individually, this criticism is not applicable.) This criticism is one that mastery learning proponents need to address. The debate, however, is one not only for educational research, but also for educational philosophy, and it will be addressed in more detail in the next section.

To summarize the research presented in this section, mastery learning has been found to work well in sequential/convergent courses, but not as well as the theory predicts. Further, some courses that are not sequential/convergent have had success with mastery learning also. There seem to be several complementary reasons for the success of mastery learning, and the research suggests a variety of payoffs for students and teachers. Critics point out, however, that mastery learning penalizes brighter students.

COMMENTS AND OBSERVATIONS

What has been attempted in this paper is a presentation of mastery learning theory: its theoretical underpinnings, operational procedures, testing considerations, and the research conducted on the theory. The presentation is admittedly somewhat biased in favor of Bloom's strategies, for two reasons. First, the author is a proponent of mastery learning. Second, it is easier and perhaps more reasonable to present an idea from the perspective of a proponent. Before closing the presentation, it is necessary to examine the criticism of mastery learning theory. We will address that criticism in this section. Following that discussion, some thoughts about where mastery learning might be headed will conclude this effort.

Criticisms of Mastery Learning Theory

In order to present the criticism of mastery learning theory, it is necessary once again to separate Bloom's ideas from Keller's ideas. But first, it should be noted that one criticism is common to both. That criticism is that the unit-to-unit mastery system only works with simplistic training-type courses. While it is probably the case that both the Bloom strategies and PSI would work well in such courses, the successes found for both systems range far beyond this type of course. It is true, however, that neither mastery learning strategy has been used widely in courses other than the sequential/

convergent type described earlier. Whether either strategy will be shown successful with less sequential, more divergent courses is a matter for speculation (and research).

The strongest criticism of mastery learning is focused on Bloom's strategies and not Keller's. This criticism is that group-based mastery learning instruction retards the growth of brighter students. When extra time is devoted to remediation of slower students, faster students are forced to sit academically idle waiting for others to catch up. This is a criticism that cannot be dismissed lightly. In order to address this issue, it is necessary to look at the educational philosophy of mastery learning.

Mastery learning theory begins with the contention that group instruction is not about to vanish from Western civilization. Group instruction is not as desirable as tutorial instruction, and no one is claiming that it ever will be. Given that it won't go away and that *some* sacrifices have to be made, mastery learning theory provides a framework which purports to minimize the sacrifices that any student will have to make. This is the standard against which mastery learning should be evaluated.

A basic position of mastery learning is that a course consists of a body of knowledge and skills and that the students in a course are essentially capable of learning the content of the course. If students are not capable of learning the course content, then they should not be in the course.

An optimal instructional strategy would be to provide each student with a skilled tutor who could take the student through the course. Unfortunately, few educational institutions have the means to provide a tutorial setting. Therefore, it is necessary to approximate that kind of setting as closely as possible.

Traditionally, the approximation has involved presenting instruction to students at a pace at which the average student (roughly speaking) can learn the material. This creates some degree of boredom after a while for the faster students and requires slower students to work harder to keep up.

PSI and other individualized systems recommend that the approximation to the tutorial setting should consist of letting the instructional materials be the tutor and having the student progress through them at his own pace. This works well for students who are motivated to work individually, but students at all levels of ability may not be motivated toward self-instruction. Also, this method places a heavy burden on the instructional materials and the teacher who has to manage the activities.

Group-based mastery learning suggests approximating the tutorial setting by presenting instruction initially in the traditional fashion, then remediating weaknesses through individualized assistance from the teacher and peer tutors. This may be of great assistance to slower students, but faster students would seem to be held back even more than under traditional instruction.

Proponents of group-based mastery learning claim that faster students are only held back in the beginning of instruction, and that after several units, mastery classes proceed through units at a faster rate than traditional instruction. In fact, in none of the research reviewed for this article was it found that *faster* students in particular learned less in mastery than nonmastery settings. The point is this: It is true that faster students cannot progress as fast as they might if they received tutorial help or if instruction was presented to the whole class at an optimal rate for the faster students, but it is contended that over the length of a course, they proceed faster than under nonmastery conditions.

Of course, proponents of PSI would contend that under PSI students can progress as quickly as they are able. That is probably true for students who work well under PSI or other individualized methods. The problem here is that many people do not learn well under those conditions.

It becomes clear that there are inherent philosophical differences among the three instructional strategies. To simplify these differences for purposes of comparison:

1. The philosophy behind PSI is that education should allow and encourage each student to learn as much as he can.
2. The philosophy behind the more traditional strategies is that education should provide opportunities for those who are willing and able to make use of them.
3. The philosophy behind Bloom's strategies is that education should ensure that all students learn those things society determines are important for them to learn.

Inherent in the PSI and traditional approaches is the notion that aptitude is fixed and that educators should accept differences in aptitude among individuals. Inherent in group-based mastery learning is the notion that the differences in aptitude among individuals are a result of our educational and societal system and that individuals can be made to be much more similar with respect to aptitude.

So, does group-based mastery learning hold faster students back? The answer from the critics remains Yes. The answer from mastery learning proponents stresses two points: (1) students are held back to an extent that is unavoidable given that educational resources do not allow tutors, and (2) slow students are a creation of our present instructional system and would not exist to the extent that they do if we modified that system.

The Future of Mastery Learning

Group-based mastery learning theory is at a fairly critical point. It is past its infancy and entering into an awkward adolescence. Part of the problem has to do with confusion about the meaning of the term. It is hoped that the real and important differences between Bloom's strategies and Keller's have been made clear here.

Part of the problem has to do with resistance to change. American education is an autonomous endeavor down to the level of the classroom teacher.* The implementation of any educational change is made more difficult by this autonomy (which is beneficial in many other respects).

Also, part of the problem is that the philosophy of mastery learning is in conflict with the philosophy of some serious and thoughtful educators. The debate, however, rarely reaches the philosophical level; it typically remains mired at the What about the faster students? level.

The future of mastery learning is still questionable. If research continues to show positive results, and if those results can start to approximate the levels claimed by the theory, then implementation would seem to be inevitable. If, on the other hand, mastery learning proves useful only in training courses and the like, then its implementation will surely and rightly be limited to those areas.

It is the belief of the author that mastery learning is the best answer available to the problems of group instruction for at least a variety of instructional areas if not all, and that the research will continue to support this view. Further, it is believed that mastery learning is not inconsistent with the view that we need exceptional individuals in order to sustain our society. Mastery learning is a theory on how to provide instruction to people in classroom settings. It assumes that their reason for being in a classroom is to learn a body of knowledge and skills. It takes the philosophical position that if all of the people are reasonably prepared to learn the course, then the charge of the educator is to teach the course to *all* of them. Not all instruction *should* take place in classroom settings with groups of people, and not all individuals *should* learn the same things.

*"Down" is used here only in a hierarchical sense. It is not meant to imply that teachers are in any perjorative sense at the bottom of the educational system.

However, no individual should be earmarked for failure at the beginning of an educational endeavor.

A Final Word

Mastery learning theory is not etched in stone. Research is still being conducted that may bring about modifications of the strategies (40). To those whose interest may have been stimulated by this discussion, it is recommended that they look at the literature mentioned here and become involved in the research.

REFERENCES*

1. Aims, D. *A Markov model for predicting performance on criterion-referenced tests.* Los Alamitos, Calif.: Southwest Regional Laboratory, June, 1971, 15 pp. ED 111 847.
2. Airasian, P.W. An application of a modified version of John Carroll's model of school learning. Unpublished master's thesis, Univer. of Chicago, 1967.
3. Airasian, P.W. Formative evaluation instruments: a construction and validation of tests to evaluate learning over short time periods. Unpublished doctoral dissertation, Univer. of Chicago, 1969.
4. Anderson, L.W. Time and school learning. Unpublished doctoral dissertation, Univer. of Chicago, 1973.
5. Annarino, A. Mastery learning in physical education. Paper presented at the annual meeting of the American Alliance for Health, Physical Education and Recreation, Milwaukee, Wisc., April, 1976, 13 pp. ED 121 772.
6. Arlin, M.N. The effects of formative evaluation on student performance. In H.F. Crombag and D.N. DeGruizter (Eds.), *Contemporary issues in educational testing.* Paris: Mouton, 1974.
7. Atkinson, R.C. Computer-based instruction in initial reading. *Proceedings of the 1967 Invitational Conference on Testing Problems.* Princeton, N.J.: Educational Testing Service, 1968. Pp. 55-67.
8. Besel, R. *A comparison of Emrick and Adam's mastery-learning test model with Kriewall's criterion-referenced test model.* (Technical Manual No. 5-71-04). Inglewood, Calif.: Southwest Regional Laboratory, 1971.
9. Block, J.H. (Ed.) *Mastery learning: theory and practice.* New York: Holt, Rinehart & Winston, 1971.
10. Block, J.H. (Ed.) *Schools, society, and mastery learning.* New York: Holt, Rinehart & Winston, 1974.
11. Block, J.H. Mastery performance standards and student learning. Unpublished study, Univer. of California, Santa Barbara, 1973.

*Items followed by an ED number (for example ED 111 847) are available from the ERIC Document Reproduction Service (EDRS). Consult the most recent issue of *Resources in Education* for the address and ordering information.

12. Block, J.H. Student evaluation: toward the setting of mastery learning performance standards. Paper presented at annual meeting of the American Educational Research Association, Chicago, Ill., 1972, 28 pp. ED 065 605.

13. Block, J.H. & Anderson, L.W. *Mastery learning in classroom instruction.* New York: Macmillan, 1975.

14. Block, J.H. & Burns, R.B. Time in school learning: an instructional psychologist's perspective. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C., 1975.

15. Block, J.H. & Tierney, M.L. An exploration of two correction procedures used in mastery learning approaches to instruction. *Journal of Educational Psychology,* December 1974, 66, No. 6, 962-967.

16. Bloom, B.S. *Human Characteristics and School Learning.* New York: McGraw-Hill, 1976.

17. Bloom, B.S. Mastery learning and its implications for curriculum development. In E.W. Eisner (Ed.), *Confronting Curriculum Reform.* Boston: Little, Brown, 1971.

18. Bloom, B.S. Learning for mastery. *Evaluation Comment.* Los Angeles: UCLA, Center for the Study of Evaluation of Instructional Programs, May 1968. ED 153 419.

19. Bloom, B.S., Hastings, J.T. & Madaus, G.F. *Handbook on formative and summative evaluation of student learning.* New York: McGraw-Hill, 1971.

20. Bosshart, Donald A. Criterion referenced achievement tests for mastery learning strategies. *Dissertation Abstracts International,* August 1973, 34 (2-A) 610.

21. Burrows, C.K. & Okey, J.R. The effects of a mastery learning strategy on achievement. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C., 1975. 13 pp. ED 109 240.

22. Carroll, J.B. A model of school learning. *Teachers College Record,* 1963, 64, 723-733.

23. Collins, K. An investigation of the variables of Bloom's mastery learning model for teaching mathematics. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Ill., 1972. 20 pp. ED 065 586.

24. Crambert, A.C. Estimation of the variables of Bloom's mastery learning model for teaching mathematics. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Ill., 1972. ED 065 589.

25. Crehan, K.D. Item analysis for teacher-made mastery tests. *Journal of Educational Measurement,* Winter 1974, 11, No. 4, 255-262.

26. Cronbach, L.J. *Essentials of psychological testing.* New York: Harper & Row, 1970.

27. Davis, W.J. The mastery learning and conventional modes of instructing college level composition: a comparative study based upon selected student characteristics. Doctoral dissertation, Oklahoma State Univer., 1975.

28. Ebel, R.L. *Essentials of psychological measurement.* Englewood Cliffs, N.J.: Prentice-Hall, 1972.

29. Emrick, J.A. An evaluation model for mastery testing. *Journal of Educational Measurement,* Winter 1971, 8, No. 4, 321-326.

30. Epstein, K.I. Predictive sample reuse and an application for criterion-referenced item analysis. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

31. Epstein, K.I. & Knerr, C.S. Criterion-referenced test interpretations of "classical" test theory. Paper presented at annual meeting of the American Educational Research Association, San Francisco, 1976. 13 pp. ED 126 154.

32. Fiel, R.L. & Okey, J.R. The effects of formative evaluation and remediation on mastery of intellectual skills. *The Journal of Educational Research*, 1974, 68, 253-255.

33. Gentile, R.J. A mastery strategy for introductory educational psychology. Unpublished materials, SUNY at Buffalo, Department of Educational Psychology, 1970.

34. Glaser, G.R. Adapting the elementary school curriculum to individual performance. *Proceedings of the 1967 Invitational Conference on Testing Problems.* Princeton, N.J.: Educational Testing Service, 1968. Pp. 3-36.

35. Glaser, G.R. Instructional technology and the measurement of learning outcomes. *American Psychologist*, 1963, 18, 519-521.

36. Hambleton, R.K. Contributions to criterion-referenced test theory: on the uses of item characteristic curves and related concepts. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

37. Hambleton, R.K. A review of testing and decision-making procedure for selected individualized instructional programs. 1972, 45 pp. ED 080 592.

38. Hapkiewicz, W.G. Mastery learning options in teaching educational psychology. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1972. 12 pp. ED 060 065.

39. Harris, C.W. An interpretation of Livingston's coefficient for criterion-referenced tests. *Journal of Educational Measurement*, 1972, 9, 27-29.

40. Hecht, L.W. *Isolation from learning supports and the processing of group instruction.* Unpublished doctoral dissertation, Univer. of Chicago, 1977.

41. Huynh, H. Two simple classes of mastery scores for domain-referenced tests. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

42. Johnston, J. & O'Neill, G. The analysis of performance criteria defining college grades as a determinate of college student academic performance. *Journal of Applied Behavior Analysis*, 1973, 6, 261-268.

43. Jones, F.G. The effects of mastery and aptitude on learning, retention, and time. April 1976, 16 pp. ED 126 381.

44. Jones, E.L., *et al.* Mastery learning: a strategy for academic success in a community college. Topical paper No. 53. December 1975, 54 pp. ED 115 315.

45. Katims, M., *et al.* The Chicago Mastery Learning Reading Program: an evaluation in an experimental setting. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

46. Keller, F.S. Goodbye, teacher. . . . *Journal of Behavior Analysis*, 1, 79-89.

47. Kersh, M.E. A strategy for mastery learning in fifth grade arithmetic. Unpublished doctoral dissertation, Univer. of Chicago, 1970.

48. King, L.A. & Szabo, M. Individualization of physics for increased enrollment through modern instructional techniques. *Science Teacher,* April 1975, 42, No. 4, pp. 36-37.

49. Kriewall, T.E. *Aspects and applications of criterion-referenced tests.* Downers Grove, Ill.: Institute for Educational Research, April 1972. 27 pp. ED 063 333.

50. Lee, Y.D., *et al.* Interaction improvement studies on the mastery learning project. (Final Report on Mastery Learning Program, 1970). Seoul, Korea: Educational Research Center, Seoul National Univer., 1971.

51. Levin, T. The effect of content-prerequisite and process-oriented experiences on application ability in the learning of probability. Unpublished doctoral dissertation, Univer. of Chicago, 1975.

52. Mayo, S.T. Measurement in education mastery learning and mastery testing. Special Report, Vol. 1, No. 3, 1970. East Lansing, Mich., March 1970. 4 pp. ED 051 299.

53. Mayo, S.T., Hunt, R.C. & Tremmel, F. A mastery approach to the evaluation of learning statistics. Paper presented at annual meeting of National Council on Measurement in Education, Chicago, Ill., 1968.

54. Meskauskas, J.A. Evaluation models for criterion-referenced testing: views regarding mastery and standard setting. *Review of Educational Research,* 1976, 46, No. 1, 133-158.

55. Millman, J. Tables for determining number of items needed on domain-referenced tests and number of students to be tested (Technical Paper No. 5). Los Angeles: Instructional Objectives Exchange, April 1972.

56. Myers, R.R. The effects of mastery and aptitude on achievement and attitude in an introductory geography course. Unpublished doctoral dissertation, Univer. of Georgia, 1975.

57. Ozcelik, D. Student involvement in the learning process. Unpublished doctoral dissertation, Univer. of Chicago, 1973.

58. Perney, J. & Huynh, H. Adaptation of the Spearman-Brown formula for the kappa reliability index in domain-referenced testing. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

59. Pillet, E.E. A study of affective and cognitive development in an introductory French course. Personal Communication. In Bloom, B.S., *Human Characteristics and School Learning.* New York: McGraw-Hill, 1976.

60. Poggio, J.P., Glassnap, D.R. & Ory, J.C. The impact of test anxiety on formative and summative exam performance in the mastery learning model. Paper presented at the annual meeting of the National Council on Measurement in Education, Washington, D.C., 1975.

61. Reichman, S.L. & Oosterhof, A.C. Strategy guidelines for the construction of mastery tests. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1976. 26 pp. ED 126 103.

62. Reynolds, C.H. & Gentile, J.R. Performance under traditional and mastery assessment procedures in relation to students' locus of control: a possible aptitude by treatment interaction. 12 pp. ED 103 357.

63. Shaycroft, M. The coefficient of accuracy: a new statistic for criterion-referenced tests. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

64. Sheldon, M.S. & Miller, E.D. Behavioral objectives and mastery learning applied to two areas of junior college instruction. Los Angeles, 1973, 90 pp. ED 082 730.

65. Skaalvik, E.M. An evaluation of mastery learning. *Scandinavian Journal of Educational Research*, 1975, 19, No. 2. Pp. 59-74.

66. Smith, J.K. Teacher planning processes and outcomes. Unpublished doctoral dissertation, Univer. of Chicago, 1977.

67. Smith, J.K. & Katims, M. Reading in the city: The Chicago Mastery Learning Reading Program. Unpublished study, New Brunswick, N.J., 1977.

68. Smith, J.K. & Wick, J.W. Practical problems of attempting to implement a mastery learning program in a large city school system. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1976. 18 pp. ED 123 563.

69. Steinhauser, F.H. A Bayesian simulation for determining mastery classification accuracy. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1976. 10 pp. ED 124 596.

70. Subkoviak, M.J. Estimating reliability from a single administration of a mastery test. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1976. 23 pp. ED 120 229.

71. Wentling, T.L. The utilization of mastery learning in vocational education: an empirical study. *Journal of Industrial Teacher Education*, Summer 1973, 10, No. 4. Pp. 36-45.

72. Wildemuth, B.M. Mastery learning and testing: an annotated ERIC bibliography. ERIC Clearinghouse on Tests, Measurement, and Evaluation, Educational Testing Service, Princeton, N.J., 1977.